

S -duality and a large N phase transition in $\mathcal{N} = 4$ on $K3$ at strong coupling

Kyriakos Papadodimas¹

*Jefferson Physical Laboratory, Harvard University
Cambridge, MA 02138, USA*

We study the supersymmetric partition function of $\mathcal{N} = 4$ super Yang-Mills with gauge group $SU(N)$ on $K3$ in the large N , fixed g limit and show that it undergoes a first order phase transition at the S -duality invariant value of the gauge coupling g . Turning on the θ -angle we find lines of phase transitions on the τ plane. The resulting phase diagram and the large N free energy are exactly $SL(2, Z)$ invariant. Similar phase transitions take place in systems related to the $\mathcal{N} = 4$ on $K3$ by dualities. One of them is the Dabholkar-Harvey heterotic string system. We consider its mixed (a la Ooguri-Strominger-Vafa) partition function allowing contributions from multi-string states. We find that in the large winding charge limit, it undergoes a phase transition with respect to chemical potential for momentum. It is a short-string, long-string transition that we find interesting in connection with black hole entropy counting.

October, 2005

¹ email: papadod@fas.harvard.edu

1. Introduction

In this paper we study the large N , fixed g limit of $\mathcal{N} = 4$ $SU(N)$ gauge theory on $K3$. Its supersymmetric partition function was computed in [1] using the fact that this theory can be twisted to a topological field theory which is simpler and in which some quantities are exactly computable. In particular the partition function of (physical) $\mathcal{N} = 4$ on $K3$ is known exactly for all values of N and g [1], [2], [3].

We show that, even though the partition function is analytic in the gauge coupling $\tau = \frac{\theta}{2\pi} + i\frac{4\pi}{g^2}$ for any finite N , it develops singularities as N goes to infinity. These correspond to large N phase transitions with respect to τ and, as it turns out, they are of first order. They are sharp transitions between a phase with no instantons at weak coupling, to a phase with a large number of instantons at strong coupling. The resulting phase diagram is $SL(2, Z)$ invariant on the τ plane, in agreement with S -duality. We find it rather interesting that even this partition function which is protected by supersymmetry and is independent of the metric of $K3$ (as it corresponds to a topological observable in the twisted theory) can undergo phase transitions.

Moreover we observe that, at least for this supersymmetric partition function, S -duality becomes more manifest in the large N , fixed τ limit. At finite N , the S -duality conjecture states [4] that $\mathcal{N} = 4$ with gauge group $SU(N)$ and coupling τ is equivalent to $\mathcal{N} = 4$ with group $SU(N)/\mathbf{Z}_N$ and coupling $-\frac{1}{\tau}$. Because the two groups are not the same, the partition function of $\mathcal{N} = 4$ is not modular invariant under the full $SL(2, Z)$. However we find that the leading piece of the large N partition function is exactly invariant under $SL(2, Z)$ indicating that S -duality may act in a simpler way in this particular large N limit.

We also consider the $\mathcal{N} = 1^*$ theory on $K3$ which is defined by a mass deformation of the $\mathcal{N} = 4$. The phase transition continues to exist after this mass deformation and so the $\mathcal{N} = 1^*$ on $K3$ undergoes a first order phase transition with respect to the microscopic coupling τ , in the large N fixed τ limit. The $\mathcal{N} = 1^*$ theory has a large number of isolated vacua in flat space. We argue that once formulated on $K3$, most of these vacua become metastable and only one of them is globally stable. For different values of the gauge coupling, different vacua become stable. So we have phase transitions between them as we vary the coupling.

The system that we study is related to various other systems that can be reached by dualities starting with Euclidean $M5$ branes wrapped on $K3 \times T^2$. The partition

functions of these systems, that count BPS states, have similar structure and we expect similar phase transitions in the large N limit, where in this case N corresponds to some charge of the system. The exact expressions for these partition functions are known and can be found in the literature. We will not repeat the analysis about the existence of a phase transition for all these systems, since the argument is almost identical with the one for the case of the $\mathcal{N} = 4$ gauge theory.

We present the equivalent phase transition in the case of the heterotic string, as we think it might be somehow related to the problem of black hole entropy counting and the recent OSV proposal [5]. We study the mixed partition function of the Dabholkar-Harvey states, with fixed magnetic (winding) charge and a chemical potential for electric charge (momentum on the S^1). We include contributions from states with more than one string. We take the large winding charge limit and find a first order phase transition with respect to the chemical potential, between a phase with many singly-wound strings and a phase with a long multi-wound string. Our calculation is not exact, but we think it captures the qualitative behavior of this system and can be made more precise if necessary. We discuss the interpretation of this phase transition in supergravity.

We briefly consider the dual picture of this transition in type IIA, corresponding to the partition function of BPS states of $D4$ and $D0$ -branes wrapped on $K3$.

The plan of this paper is as follows: in section 2 we discuss the relation between the physical $\mathcal{N} = 4$ gauge theory and its topological version. In section 3 we present the partition function of $\mathcal{N} = 4$ on $K3$. In section 4 we study the large N limit of this partition function, we identify its singularities and we discuss the interpretation of the phase transitions in terms of instantons in the gauge theory. In section 5 we study the same transition in the $\mathcal{N} = 1^*$ theory on $K3$. In section 6 we turn on the θ angle and describe the $SL(2, Z)$ invariant phase diagram of the $\mathcal{N} = 4$ on $K3$. In section 7 we relate this with similar transitions in systems connected by string dualities. In section 8 we discuss the heterotic Dabholkar-Harvey string and in 9 the $D4 - D0$ system. In sections 10 and 11 we conclude with some questions that we find interesting.

2. Topological and Physical $\mathcal{N} = 4$

2.1. The topological theory

In this paper we will be mostly interested in the physical $\mathcal{N} = 4$. However the topological version of the same theory is useful for the computation of some quantities

that are protected by supersymmetry. For more details on the topological $\mathcal{N} = 4$ see [6] and references therein.

It is well known that $\mathcal{N} = 4$, $SU(N)$ SYM on a 4-manifold M can be twisted to a topological theory [7], [1], [8]. This is achieved by turning on background gauge fields for the $SU(4)$ R -symmetry, equal to the spin connection of the manifold². As a result of this, two of the supercharges become scalars and remain unbroken even when the theory is defined on a curved manifold. The observables that are in the cohomology of these supercharges are topological. In particular the partition function for these models is a topological invariant³.

For the topological $\mathcal{N} = 4$ theories (and unlike the topological $\mathcal{N} = 2$ [10]) the partition function *does* depend on the gauge coupling $\tau = \frac{\theta}{2\pi} + i\frac{4\pi}{g^2}$. This property was used in [1] to test S -duality for $\mathcal{N} = 4$ at strong coupling. In the same paper it was shown that (at least when the theory is defined on certain manifolds, including the case of $K3$) the partition function is the generating function for the Euler characteristic of instanton moduli spaces for $SU(N)$ gauge theory on M . The partition function has the general form:

$$Z_N(\tau) \sim q^{-\frac{N\chi}{24}} \sum_{k=0}^{\infty} c_{k,N} q^k, \quad q = \exp(2\pi i\tau) \quad (2.1)$$

where $c_{k,N}$ is the Euler characteristic of the moduli space of k -instantons of $SU(N)$ on M , χ is the Euler characteristic of the manifold M . For $K3$ we have $\chi = 24$.

2.2. Relation to Physical $\mathcal{N} = 4$

On a general 4-manifold the twisted and the physical theories are different. However for hyper-Kähler manifolds the twisted and the physical $\mathcal{N} = 4$ coincide [1]. This means

² For the $\mathcal{N} = 4$ theory there are 3 different twistings that lead to topological theories, depending on how we embed the rotation group $SO(4) \sim SU(2)_L \times SU(2)_R$ into the $SU(4)$ R -symmetry group. Here we are interested in the twisting considered in [1].

³ One way to compute this partition function for Kähler manifolds is to introduce a mass perturbation for the chiral multiplets of $\mathcal{N} = 4$ that does not spoil the topological nature of the theory, but which breaks supersymmetry down to $\mathcal{N} = 1$. This $\mathcal{N} = 1$ theory has a mass gap (more about it in section 5). Using the topological invariance of the theory we can send the size of the manifold to infinity. Then because the theory is gapped, the partition function gets contributions only from the ground states of this $\mathcal{N} = 1$ theory [1], [2]. An analogous method was used earlier in [9] for the computation of the Donaldson invariants in the topological $\mathcal{N} = 2$ theory [10].

that the (supersymmetric) partition function of the *physical* $\mathcal{N} = 4$ is equal to the partition function of the twisted theory on the same manifold, which as we said above is a topological and, in some cases, exactly computable function for any N and g .

The two compact 4-dimensional hyper-Kähler manifolds are T^4 and $K3$. For T^4 the partition function is constant, so there is not much we can say about it. On the other hand, the topological partition function for $K3$ is nontrivial. Since $K3$ is hyper-Kähler, it is the same as the partition function of the physical $\mathcal{N} = 4$ on $K3$.

3. $\mathcal{N} = 4$ super Yang-Mills on $K3$

3.1. The Vafa-Witten partition function

In [1] Vafa and Witten computed the partition function of $SU(N)$ $\mathcal{N} = 4$ on $K3$, for N prime. Their result was extended to non prime N in [2]. See also [3].

For the gauge group $SU(N)$ with N prime, the partition function takes the form:

$$Z(\tau) = \frac{1}{N^3} G(N\tau) + \frac{1}{N^2} \sum_{m=0}^{N-1} G\left(\frac{\tau+m}{N}\right), \quad (3.1)$$

where

$$G(\tau) = \eta(\tau)^{-24}, \quad (3.2)$$

$\eta(\tau)$ is the Dedekind eta function and $\tau = \frac{\theta}{2\pi} + i\frac{4\pi}{g^2}$ the complexified gauge coupling.

For non-prime N the partition function is somewhat more complicated and is almost a Hecke transformation of order N of the function $G(\tau)$:

$$Z(\tau) = \frac{1}{N^3} \sum_{0 \leq a, b, d \in \mathbb{Z}; ad=N; b \leq d-1} d G\left(\frac{a\tau+b}{d}\right). \quad (3.3)$$

Clearly for N prime (3.3) reduces to (3.1).

Note that $G(\tau)$ is the same⁴ as the partition function of the Dabholkar-Harvey BPS heterotic string states [11], [12], [13].

One way to understand the appearance of the heterotic string partition function is the following [2] : we consider the supersymmetric partition function of N $M5$ branes on $K3 \times T^2$. We can dimensionally reduce on the T^2 and we end up with $\mathcal{N} = 4$ on $K3$, whose partition function we are studying⁵. The same partition function can be computed in the following way: the N $M5$ branes can be thought of as N $NS5$ branes in IIA on $K3 \times T^2$ [14]. Then by IIA-Heterotic duality we end up with the heterotic string wound N times around the T^2 (all these in the Euclidean theories).

⁴ Up to an overall factor of 16.

⁵ Up to some extra factors explained in [2].

3.2. Some comments on S -duality

What is the behavior of $Z(\tau)$ under S -duality transformations?

The original S -duality conjecture for the $\mathcal{N} = 4$ super Yang-Mills states that the theory with gauge group G and coupling τ is equivalent to the theory with gauge group \hat{G} (the dual of G)⁶ and coupling $-\frac{1}{\tau}$. The dual group of $SU(N)$ is $SU(N)/\mathbf{Z}_N$, which is not simply connected and admits 't Hooft magnetic fluxes (corresponding to gauge bundles with fractional instanton number), unlike the $SU(N)$ bundles for which the instanton number is integer. Therefore when we compute the partition function of $SU(N)/\mathbf{Z}_N$ we have to sum over these flux sectors. This means that in general $Z_{SU(N)}(\tau) \neq Z_{SU(N)/\mathbf{Z}_N}(\tau)$. What we expect from S -duality is that :

$$\begin{aligned} Z_{SU(N)}(\tau + 1) &\sim Z_{SU(N)}(\tau) \\ Z_{SU(N)}\left(-\frac{1}{\tau}\right) &\sim Z_{SU(N)/\mathbf{Z}_N}(\tau). \end{aligned} \tag{3.4}$$

So the partition function of $SU(N)$ is not exactly invariant under $SL(2, \mathbb{Z})$ transformations. Moreover even (3.4) has to be somewhat modified by assigning a nonzero modular weight to the function Z , once the theory is formulated on a curved manifold. All these issues have been analyzed in detail in [1]. See also [15], [16].

However, since we know the *exact* partition function of $\mathcal{N} = 4$ on $K3$ for all τ and N we can forget about these subtleties related to S -duality and study how the function (3.3) transforms under $SL(2, \mathbb{Z})$ transformations of the variable τ (without changing the group with the dual group).

First, from the modular transformations of the function $G(\tau) = \eta(\tau)^{-24}$ we find that the partition function (3.1), (3.3) is invariant under $\tau \rightarrow \tau + 1$:

$$Z_{SU(N)}(\tau + 1) = Z_{SU(N)}(\tau). \tag{3.5}$$

Second, one can show (see [3] for details) that:

$$Z_{SU(N)}\left(-\frac{1}{\tau}\right) = (N\tau)^{-12} \frac{1}{N^2} \sum_{a,b,d;p=gd\{b,d\}} d^{12} p^{11} G\left(\frac{a\tau + b}{d}\right) \tag{3.6}$$

We would like to emphasize that in (3.5), (3.6) we are just looking at the way (3.1), (3.3) transform under $SL(2, \mathbb{Z})$ transformations of the variable τ . These are not the actual

⁶ The weight lattice of \hat{G} is the dual of that of G .

S -duality transformations, as we are not exchanging the group $SU(N)$ with $SU(N)/\mathbf{Z}_N$. We see that Z is not exactly invariant under these naive $SL(2, Z)$ transformations of τ , which of course is not surprising as the S -duality conjecture requires the exchange of the gauge group with its dual group and thus the inclusion of magnetic flux sectors.

In this paper we are interested in the leading term of the partition function of the $SU(N)$ gauge theory in the large N limit. In the next section we analyze it and find that it is exactly $SL(2, Z)$ invariant. Thus S -duality may act in a simpler way in the this large N , fixed τ limit.

4. Large N limit and Phase Transitions

4.1. The Phase Transition

For finite N the $SU(N)$ Vafa-Witten partition function is analytic in τ . However it is possible that in the large N limit the partition function develops singularities, and we can have large N phase transitions even when the theory is defined on compact volume (an early example is [17]). In our system of $\mathcal{N} = 4$ on $K3$ we will see that in the large N limit the partition function takes the form:

$$Z_{SU(N)}(\tau) \sim \exp(Nf_\infty(\tau)) + \text{subleading}. \quad (4.1)$$

For any finite N the free energy $f_N(\tau) = \frac{\log Z_N(\tau)}{N}$ is an analytic function of τ . On the other hand, the large N free energy $f_\infty(\tau) = \lim_{N \rightarrow \infty} f_N(\tau)$ is continuous but not everywhere smooth. This means that the system undergoes large N phase transitions as we vary the gauge coupling τ .⁷

For simplicity let us start with the case where the θ -angle is zero, so $\tau = i\frac{4\pi}{g^2}$ is purely imaginary. In appendix B we show that for $\theta = 0$, there are only two terms that can

⁷ We can understand how this phase transition is possible from the following observation. The partition function (3.1) or (3.3) is a sum of many terms. The argument of the function G of some of these terms goes either to infinity or to the real axis in the limit under consideration. Using the modular properties of G and its asymptotic expansion (see Appendix A) it is easy to show that some of these terms grow like $e^{a_i(\tau)N}$ in the large N limit. The function $a_i(\tau)$ is different for each term. From this we conclude that the partition function in the large N limit will be dominated by the term with the largest $Re(a_i(\tau))$. Moreover as we vary τ it is possible that different terms (with different functions $a_i(\tau)$) become dominant. Thus, while each $a_i(\tau)$ are analytic, the resulting $f_\infty(\tau)$ may have non-analytic behavior.

compete for dominating the partition function in the large N limit for various values of g . These are the terms $G(N\tau)$ and $G(\frac{\tau}{N})$. Their asymptotic behavior for large N is $\exp\left(N\frac{8\pi^2}{g^2}\right)$ and $\exp\left(N\frac{g^2}{2}\right)$ respectively. So in the large N limit only one of the two will dominate. Comparing them we conclude that:

- At weak coupling ($g < \sqrt{4\pi}$):

$$Z(\tau) \sim G(N\tau) \sim \exp\left(N\frac{8\pi^2}{g^2}\right) + \dots, \quad (4.2)$$

- at strong coupling ($g > \sqrt{4\pi}$):

$$Z(\tau) \sim G\left(\frac{\tau}{N}\right) \sim \exp\left(N\frac{g^2}{2}\right) + \dots \quad (4.3)$$

The large N free energy $f_\infty(g)$ is ⁸:

$$f_\infty(g) = \begin{cases} \frac{8\pi^2}{g^2} & \text{if } g < \sqrt{4\pi}; \\ \frac{g^2}{2} & \text{if } g > \sqrt{4\pi}. \end{cases}$$

So, the free energy is continuous, but its first derivative is discontinuous at the value $g_c = \sqrt{4\pi}$, precisely the S -duality invariant value of the coupling. The theory undergoes a first order phase transition at this value of the coupling.

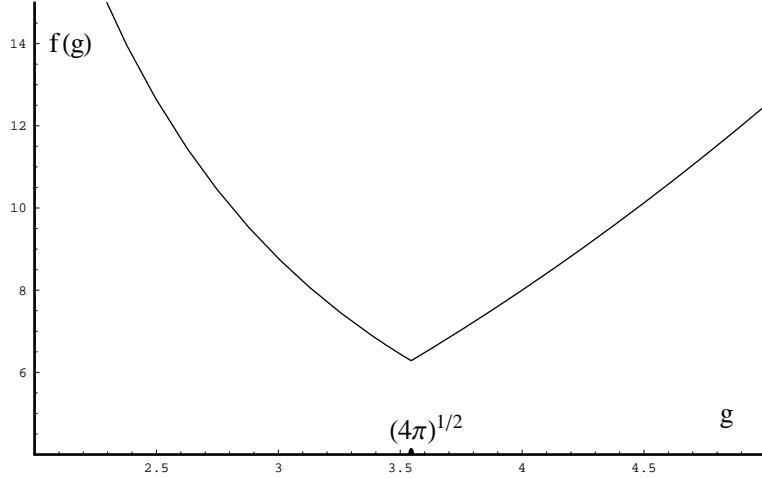


Fig. 1: Large N free energy of $\mathcal{N} = 4$ on $K3$ as a function of the gauge coupling g for $\theta = 0$.

⁸ In this section we are writing f_∞ as a function of g instead of τ since we have assumed $\theta = 0$.

4.2. Physical Interpretation in terms of Instantons

What is the interpretation of this phase transition?

According to the analysis of [1] the partition function of $\mathcal{N} = 4$ on $K3$ has the form :

$$Z_N(\tau) = q^{-N} \sum_{k=0}^{\infty} c_{k,N} q^k \quad (4.4)$$

where $q = e^{-\frac{8\pi^2}{g^2}}$ and $c_{k,N}$ is the Euler characteristic of the k -instanton moduli space for $SU(N)$ on $K3$.

We see that at weak coupling the partition function goes like $\exp\left(\frac{N8\pi^2}{g^2}\right)$. This means that the partition function is dominated by the zero instanton sector at weak coupling.

At strong coupling the partition function goes like $\exp\left(N\frac{g^2}{2}\right)$. At first sight this does not look like an instanton configuration as the dependence on g is of different form from the usual $\exp\left(-k_o\frac{8\pi^2}{g^2}\right)$ for some number k_o of instantons. However it is possible that the dominant configuration has instanton number that goes like $k_o \sim Ng^4$ for large g , explaining the g -dependence⁹. If we assume that at strong coupling the partition function is dominated by a certain k -instanton sector with k large, we can use an asymptotic form for $c_{k,N}$ and try to evaluate the partition function using the saddle point method. We assume that the Euler characteristic of the k instanton moduli space for $SU(N)$ on $K3$ goes like $\exp(4\pi\sqrt{kN - N^2})$ for large N and k and $\frac{k}{N} \sim \text{const} > 1$ (see Appendix C for some details and also [18], [19]). We have¹⁰ :

$$\begin{aligned} Z &\sim \exp\left(N\frac{8\pi^2}{g^2}\right) \sum_{k=N}^{\infty} \exp(4\pi\sqrt{kN - N^2}) \exp\left(-k\frac{8\pi^2}{g^2}\right) \\ &\sim \exp\left(N\frac{8\pi^2}{g^2}\right) \int_{x=1}^{\infty} \exp(N4\pi\sqrt{x-1}) \exp\left(-Nx\frac{8\pi^2}{g^2}\right) dx, \quad x = \frac{k}{N} \end{aligned} \quad (4.6)$$

In the large N limit we find the saddle point for the integral:

⁹ I would like to thank S.Minwalla for very useful comments and suggestions .

¹⁰ Ignoring the contribution from the zero instanton sector which is exponentially suppressed at strong coupling. Also the dimension of the k -instanton moduli space for $SU(N)$ on $K3$ is [1] :

$$\dim \mathcal{M}_k^N = 4kN - 4(N^2 - 1) \quad (4.5)$$

which is negative for $1 \leq k < N$. This means that instanton sectors start at $k \geq N$ after the trivial configuration $k = 0$.

$$x_{sp} = \frac{g^4}{(4\pi)^2} + 1, \quad \text{or} \quad , \quad k_{sp} = \frac{g^4}{(4\pi)^2} N + N \quad (4.7)$$

and the value of the integral is :

$$Z_{sp} \sim \exp \left(N \frac{g^2}{2} \right) \quad (4.8)$$

which is what we found before from the exact the partition function at strong coupling¹¹. To summarize we see that a possible explanation for this phase transition is that for all values of the coupling there are two saddle-points/phases that are competing. One is the zero instanton phase and the other a phase with a very large number of instantons given by (4.7). At strong coupling the system is dominated by the many-instanton phase. As we decrease the coupling the dominant instanton number decreases and at some point the phase with no instantons has lower free energy and we have a sharp first order phase transition to the zero instanton sector. Figure 2 shows the free energy diagram of the two competing phases for different values of the coupling, and the transition at $g = \sqrt{4\pi}$.

We can also compute the entropy of the two phases, if we assume that the role of "temperature" is played by g^2 :

- For weak coupling¹²(low temperature):

$$S_w = 0, \quad E_w = -8\pi^2 N, \quad F_w = -8\pi^2 N \quad (4.9)$$

- For strong coupling (high temperature):

$$S_s = g^2 N, \quad E_s = \frac{g^4}{2} N, \quad F_s = -\frac{g^4}{2} N \quad (4.10)$$

Our conclusion is that the phase transition is between a weak coupling phase of no instantons with small entropy and low energy, and a strong coupling phase with many instantons, higher energy and entropy.

¹¹ We notice that the saddle point solution gives $k_{sp} > N$, so we are in an instanton sector with a moduli space of positive dimension. Also our approximation for $c_{k,N}$ is justified since k, N are large and $k/N \sim \text{const} > 1$.

¹² In this section by free energy we mean: $F = -g^2 \text{Log}(Z)$, the thermodynamic free energy, if we consider g^2 as temperature. In the other sections of the paper we use the term free energy for $f = \frac{\text{Log} Z}{N}$. We hope that this change does not cause confusion.

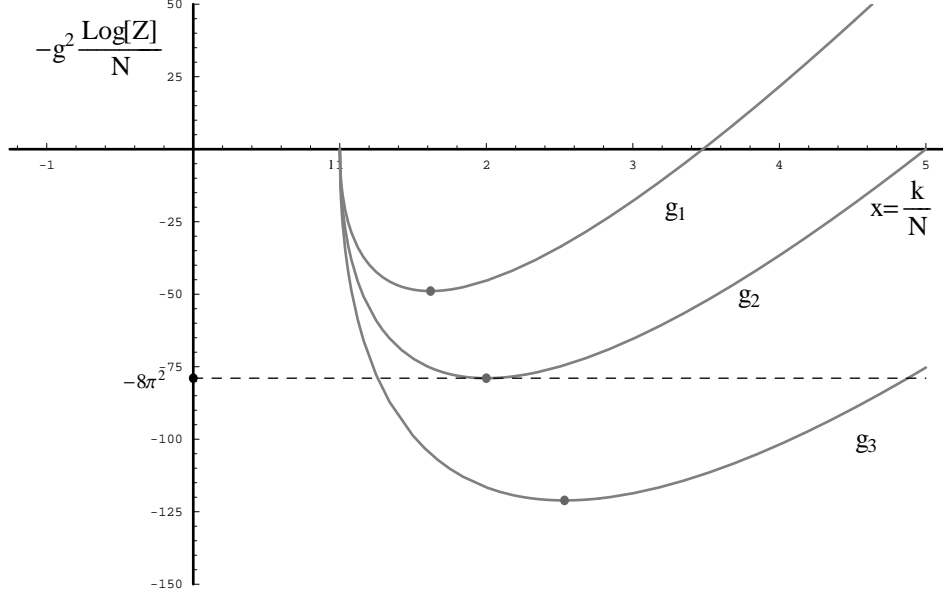


Fig. 2: Free energy as a function of the dominant instanton number k for three different values of the gauge coupling $g_1 < g_2 = \sqrt{4\pi} < g_3$. There are always two phases that are competing. One is the isolated "saddle point" at $x = 0$ (no instantons), whose free energy does not depend on g . The other is the minimum of the curves shown, corresponding to the saddle point described in the text. As we increase the gauge coupling g this saddle point becomes "thermodynamically" favored and at g_2 we have a first order phase transition between the two phases.

5. Mass deformation and the phase transition for $\mathcal{N} = 1^*$

It is known that the $\mathcal{N} = 4$ theory can be deformed to the $\mathcal{N} = 1^*$ theory by giving a mass term to its chiral multiplets [1], [2], [20]. We argue that the phase transition we are studying exists also in this $\mathcal{N} = 1^*$ theory on $K3$ in the same large N , fixed g limit. In the following section we briefly describe some basic properties of the $\mathcal{N} = 1^*$ theory and then we discuss the characteristics of the phase transition in $\mathcal{N} = 1^*$.

5.1. The $\mathcal{N} = 1^*$ theory

The $\mathcal{N} = 4$ super Yang-Mills in $\mathcal{N} = 1$ language contains a vector multiplet and 3 chiral multiplets $\Phi_i, i = 1, 2, 3$ in the adjoint of the gauge group, and a superpotential:

$$W = Tr([\Phi_1, \Phi_2]\Phi_3) \quad (5.1)$$

We add mass terms for the chiral multiplets to the superpotential:

$$\Delta W = \frac{1}{2} m \text{Tr} (\Phi_1^2 + \Phi_2^2 + \Phi_3^2) \quad (5.2)$$

The effect of this term is to break supersymmetry down to $\mathcal{N} = 1$. This theory is called $\mathcal{N} = 1^*$. It does not have a moduli space, but a number of isolated vacua, which have been studied in [1], [2], [20].

The perturbation (5.2) does not spoil the topological nature of the theory (as it is BRST exact) on Kähler manifolds, and there is one unbroken supercharge even after the mass perturbation. This means that even if we turn on this mass deformation in the $\mathcal{N} = 4$ theory, its partition function is independent of m . So the expression (3.3) is the (supersymmetric) partition function of $\mathcal{N} = 1^*$ $SU(N)$ on $K3$ for any value of m and any size of the $K3$.

Let us introduce a parameter t that characterizes the overall size of the $K3$. That is, starting with an arbitrary metric $g_{\mu\nu}$ on $K3$ we parametrize different metrics by $t^2 g_{\mu\nu}$, controlling the size of the manifold. The partition function (3.3) is independent of m and t , and exhibits a large N phase transition with respect to the gauge coupling τ for all values of m and t .

We consider the parameter tm . This parameter smoothly interpolates between the original $\mathcal{N} = 4$ at small tm and the $\mathcal{N} = 1^*$ *in its vacuum state* at large tm . The phase transition with respect to τ exists for all values of these parameters but its interpretation is different in the two limits mentioned above.

At small tm we effectively have the $\mathcal{N} = 4$ on $K3$ and as we saw in the previous section the phase transition is interpreted as a transition between a configuration with no instantons at weak coupling to a phase with many instantons at strong coupling.

What about the limit of large tm ?

As explained in [9], [1], [2], we expect that in this limit the partition function will be determined by the (flat space) vacuum states of the $\mathcal{N} = 1^*$ theory. The vacuum structure of $\mathcal{N} = 1^*$ has been analyzed in [1], [2], [20]. Classically the vacua are given by the solutions of:

$$[\Phi_i, \Phi_j] = m \epsilon_{ijk} \Phi_k \quad (5.3)$$

These equations are the commutation relations of $SU(2)$. So the classical vacua of the theory are characterized by an N -dimensional (possibly reducible) representation of $SU(2)$.

Quantum mechanically there are two categories of vacua, the massive vacua where there are no unbroken $U(1)$ factors and there is a mass gap, and the massless vacua. The massless vacua with the $U(1)$ factors have extra fermion zero modes and do not contribute [2] to the supersymmetric partition function that we are studying, so we only need to consider the vacua that have a mass gap¹³.

The two extreme cases are:

- i. The $SU(N)$ confined vacuum: here $\Phi_i = 0$, and the unbroken group is $SU(N)$. Quantum mechanically the theory confines, a mass gap is generated, chiral symmetry is broken and the classical vacuum splits into N quantum vacua with different expectation values for the gluino condensate. These N vacua are cyclically permuted under $\tau \rightarrow \tau + 1$. The contribution to the partition function of $\mathcal{N} = 4$ on $K3$ from these vacua is given by the terms:

$$\text{confined } SU(N) \text{ vacuum : } \quad \frac{1}{N^2} \sum_{m=0}^{N-1} G\left(\frac{\tau+m}{N}\right) \quad (5.4)$$

in (3.3).

- ii. The other extreme case is when the $SU(N)$ group is fully higgsed. This happens when the expectation values of the scalars transform in the N -dimensional irreducible representation of $SU(2)$. Then a mass gap is generated even classically. The contribution to the partition function from this vacuum is:

$$\text{fully higgsed vacuum : } \quad \frac{1}{N^3} G(N\tau) \quad (5.5)$$

There are other intermediate *massive* vacua when N has divisors. If for example $N = ad$, there is a vacuum where the expectation values of the Φ 's transform in d copies of the a -dimensional irreducible representation of $SU(2)$. Then the unbroken gauge group is $SU(d)$. Quantum mechanically it confines and splits into d vacua. The contribution from these vacua is:

$$\text{intermediate vacua : } \quad \frac{1}{N^3} \sum_{b=0, \dots, d-1} dG\left(\frac{a\tau+b}{d}\right) \quad (5.6)$$

So as explained in [2] we see that all massive vacua of the $\mathcal{N} = 1^*$ theory are in one to one correspondence with the various terms of the Hecke transformation that give the partition function (3.3).

All these vacua are interchanged by the action of S -duality in $\mathcal{N} = 4$.

¹³ In the limit $N \rightarrow \infty$ with g constant, the mass gap is of the order of the mass perturbation m [20].

5.2. Phase transition in $\mathcal{N} = 1^*$

We saw above that in flat space the $\mathcal{N} = 1^*$ theory has a number of (supersymmetric) massive vacua, all of course with zero energy.

Now let us consider the same theory on compact volume, starting with the example of a flat T^4 . Let us choose a small region of the T^4 and look at the fields in that region. We can send t to infinity¹⁴, blowing up the size of this region as much as we want without changing the partition function. For very large t and because the theory has a mass gap¹⁵ we expect to find the fields in this region in their (flat space) vacuum configuration. But which of the many vacua of $\mathcal{N} = 1^*$? Since we are in compact volume we can't have spontaneous symmetry breaking (at finite N). The wavefunction of the system is not localized around one vacuum only, but is in superposition of the various vacua.

Now consider the theory on $K3$. As in the case of T^4 the partition function receives contributions from all (flat space) vacua as we can see from the expression (3.3). However there is a difference. We notice that because of curvature corrections the contributions from different vacua are not equal and depend on the coupling constant τ .¹⁶ Notice that these contributions are not extensive in the volume of $K3$ (as they are independent of t), and they only depend on its topology. So once we are on $K3$ these curvature corrections make most of the (flat space) vacua become metastable and precisely which one is globally stable depends on the value of the coupling τ . But since we are in finite volume and at finite N we still get contributions from all of them¹⁷, and this is reflected in the form of (3.3).

In the large N limit spontaneous symmetry breaking can take place even in finite volume. Even if there is a very small difference in "free energy" between two metastable

¹⁴ Again, t parametrizes the overall size of the manifold.

¹⁵ As explained in [2] vacua without mass gap do not contribute to the partition function because of extra fermionic zero modes.

¹⁶ These are the gravitational corrections that correspond to the genus 1 diagrams in the Dijkgraaf-Vafa matrix model around the corresponding vacuum of $\mathcal{N} = 1^*$ [21].

¹⁷ As in any quantum system in compact volume with metastable vacuum states. In our system the energy difference between the vacua is not proportional to the volume t of the $K3$, so when we send t to infinity we do not have a decay to the stable vacuum. On the other hand the energy difference is proportional to N so in the large N limit we have the localization of the wavefunction of the system around the absolutely stable vacuum, and phase transitions as we change the coupling.

vacua, if this difference scales like N then in the large N limit the theory jumps to the vacuum with the smaller free energy, depending on the value of the coupling. The contributions from other vacua are exponentially suppressed. For example for $\theta = 0$: at weak coupling g the most stable vacuum is the fully higgsed (corresponding to the term $G(N\tau)$). At strong coupling the confined $SU(N)$ vacuum (corresponding to $G(\frac{\tau}{N})$) becomes more stable and we have a first order phase transition between the two at $g = \sqrt{4\pi}$.

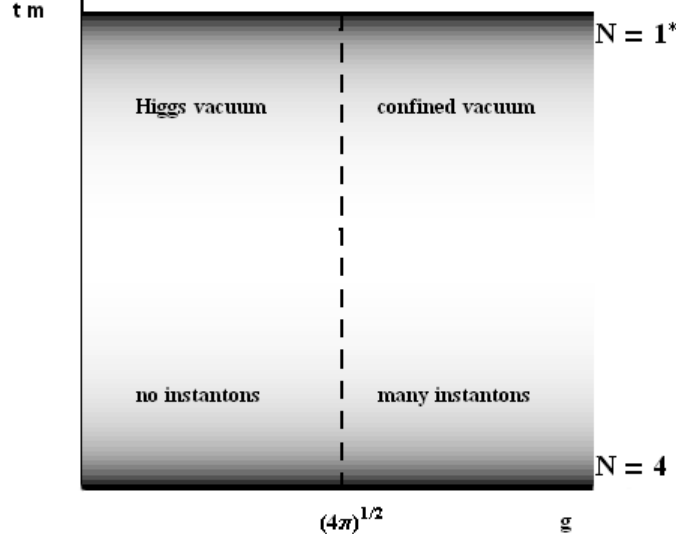


Fig. 3: Phase Diagram of $\mathcal{N} = 1^*$ $SU(N)$ on $K3$. We see the phase transition line at $g = \sqrt{4\pi}$ for all values of tm . At large tm the most natural physical interpretation of the phase transition is as a first order transition between two different "ground states" of $\mathcal{N} = 1^*$ on $K3$. At small tm the theory looks like $\mathcal{N} = 4$ on $K3$ and as we concluded before the phase transition is between a phase with no instantons to a phase with many.

To summarize, for any value of the coupling there are many metastable vacua for $\mathcal{N} = 1^*$ on $K3$. At infinite N the theory selects the most stable vacuum, but which one it is, depends on the value of the coupling. As we change the coupling different competing metastable vacua become absolutely stable and we have first order phase transitions between them. At finite N the theory is in a superposition of all vacua, as we can see from the (3.3).

In figure 3 we see the phase diagram as a function of the gauge coupling g and the parameter tm that controls the amount of breaking of the $\mathcal{N} = 4$ supersymmetry. In the two limits of large and small tm the interpretation of the phase transition is more clear.

For intermediate values of tm the phase transition still happens but is more difficult to interpret physically.

6. Turning on the θ angle and an $SL(2, Z)$ invariant phase diagram

We now turn to the case of non-zero θ . Again for every value of the coupling τ we have to identify the term that dominates the partition function, compute the large N free energy $f_\infty(\tau)$ and study its singularities.

This can be easily done by using the modular properties (3.5) and (3.6) of the function $Z_{SU(N)}(\tau)$. We make the following observation: we know (see Appendix B) that for any τ in the first fundamental domain of $SL(2, Z)$:

$$\mathcal{F}_o = \{|\tau| > 1, \quad |Re(\tau)| < \frac{1}{2}, \quad Im(\tau) > 0\} \quad (6.1)$$

the large N partition function is dominated by the term $G(N\tau)$, while all others are (exponentially) suppressed compared to this largest term. From this observation and (3.5), (3.6), we conclude that for two different values of τ which are related by an $SL(2, Z)$ transformation, the dominant term has the same value or more precisely the same exponential growth¹⁸ with N .

This has two interesting consequences. First, it means that we only need to study the large N limit of the partition function in the first fundamental domain \mathcal{F}_o and then extend it everywhere else on the plane by $SL(2, Z)$ transformations. Second, the same property means that in the large N limit the free energy becomes exactly invariant under S -duality transformations without the exchange of the gauge group with its dual group.

In the first fundamental domain it is always the term $G(N\tau)$ that dominates, giving the free energy:

$$f_\infty(\tau) = -2\pi i\tau, \quad \tau \in \mathcal{F}_o \quad (6.2)$$

and from our arguments above the function must be extended everywhere else on the plane by $SL(2, Z)$ transformations, as if it were a modular form of weight 0. That is:

$$f_\infty(\tau) = -2\pi i \frac{a\tau + b}{c\tau + d} \quad (6.3)$$

¹⁸ The prefactor of the exponential will be different in general, but it is at most polynomial in N and drops out when we compute the large N free energy.

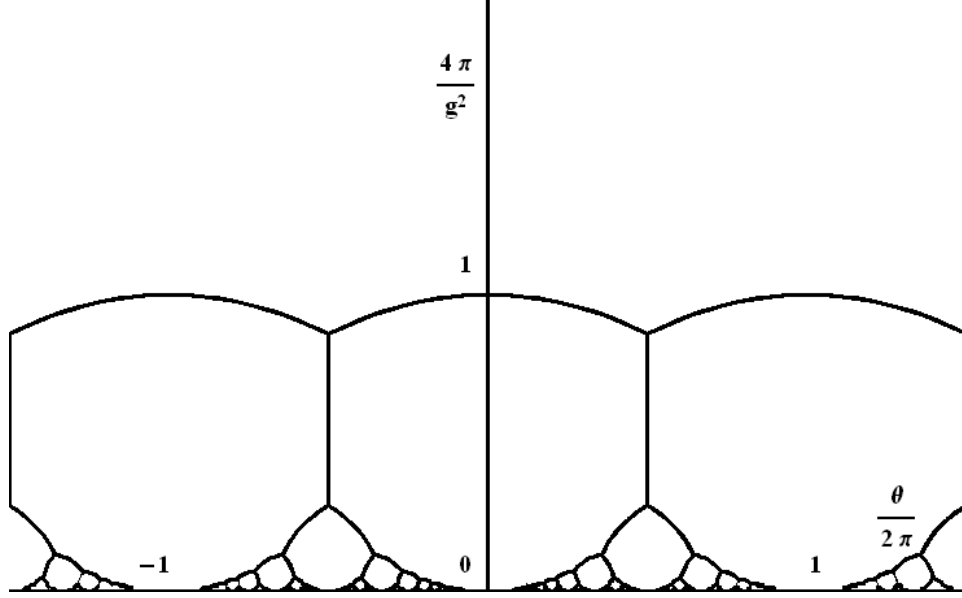


Fig. 4: Phase Diagram of $\mathcal{N} = 4$ $SU(N)$ on $K3$, in the large N fixed τ limit. The lines correspond to first order phase transitions.

where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is the unique $SL(2, Z)$ transformation that maps any given τ , inside the first fundamental domain \mathcal{F}_o .

So the conclusion is that the large N free energy transforms like a modular form of weight 0 under $SL(2, Z)$ transformations. Thus we see that S -duality acts in a simpler way in the theory in the large N limit.

However, as we saw in the previous section the free energy it is not necessarily analytic. It is easy to see that this large N free energy $f_\infty(\tau)$ is continuous everywhere¹⁹, but its first derivative with respect to the coupling is discontinuous along the lower boundary of the first fundamental domain $|\tau| = 1$, $|Re(\tau)| < \frac{1}{2}$, as well as on its images under all $SL(2, Z)$ transformations. So we have first order phase transitions with respect to the gauge coupling τ along these lines. In figure 4 we can see these lines of first order phase transitions, that characterize the $SL(2, Z)$ invariant phase diagram for $\mathcal{N} = 4$ on $K3$. For a similar phase diagram see [22], [23].

¹⁹ There are some trivial discontinuities along the vertical boundaries of the first fundamental domain \mathcal{F}_o and all its images under $SL(2, Z)$, on which the imaginary part of the free energy jumps by 2π . Of course these are not real physical singularities as the partition function $Z \sim \exp(Nf_\infty)$ is continuous under this jump of the imaginary part of f_∞ .

7. Connections with other systems

There are BPS systems of D-branes, related by dualities, whose supersymmetric partition function has a similar structure: it is an order N Hecke transformation of some smooth function of τ . N corresponds to some charge of the system and τ to a chemical potential or complex modulus of a T^2 in spacetime. The partition function is a smooth function of τ for finite N but it develops singularities as N goes to infinity. These systems will exhibit similar first order phase transitions with respect to τ . We mention some of them:

- N (Euclidean) $M5$ branes in M theory wrapped on $K3 \times T^2$. This system was discussed in [2]. Its partition function is an (almost) Hecke transformation of order N of some smooth modular form. In the large N limit we expect a phase transition with respect to the complex modulus τ of the T^2 , similar to the one we studied in this paper.
- Similarly for N $NS5$ branes in type IIA wrapped on $K3 \times T^2$. This system was studied in [23] and the story is similar with the cases above.
- The low energy world volume theory of these N $M5$ branes is the $6D$ theory A_{N-1} SCFT on $K3 \times T^2$. These theories (on flat space) have no parameters. From the above we conclude that if we formulate the A_{N-1} theory on $K3 \times T^2$ and send N to infinity, we expect a first order phase transition as a function of the modular parameter τ of the T^2 .
- In type IIA the partition function for BPS states of N $D4$ branes wrapped on $K3$ and with chemical potential τ for $D0$ brane charge and by T -duality, in IIB the partition function for BPS states of $D1 - D5$. See also [22], [23].
- Finally as explained in [2], by IIA-Heterotic duality we have the same structure for the partition function which counts BPS states for the heterotic string wound N times on a T^2 . (Note that this is in Euclidean signature, so we must be careful about the interpretation of the multi-wound heterotic string). The partition function depends on the complex parameter τ of the T^2 . In the large N limit we find first order phase transitions with respect to τ . The different phases correspond to configurations where the heterotic string wraps a specific cycle of the T^2 N times. As we change τ this distinguished cycle changes, making the large N partition function non-smooth.

8. Heterotic String Transition

We find the case of the multi wound heterotic string rather interesting in connection with black hole entropy counting and the Ooguri-Strominger-Vafa proposal [5]. According to the OSV prescription the partition function of the black hole has to be computed with fixed magnetic charge, arbitrary electric charge and with some chemical potential for electric charge. Motivated by the existence of the phase transition studied above, we will consider the partition function of the Dabholkar-Harvey states of the heterotic string with fixed magnetic charge W (winding) and fixed chemical potential ϕ_0 for the electric charge P (momentum on the S^1). We will include contributions from states that consist of more than one strings. We find a phase transition with respect to the chemical potential ϕ_0 in the large W limit.

It is well known that the heterotic string compactified on an S^1 has an infinite tower of massive BPS states (Dabholkar-Harvey states). These are states of winding w and momentum p . The right moving (supersymmetric) sector of the heterotic string is in its ground state, while the left moving sector can be excited to an arbitrary oscillator level N_L . The only constraint is the level matching condition:

$$N_L = 1 - pw \quad (8.1)$$

All of these states are BPS [11], [12], [13].

We want to compute the partition function of these states with fixed total winding charge W and chemical potential ϕ_0 for total momentum P on the S^1 , *allowing contributions from multi-string states*²⁰. So we want to compute:

$$Z_W(\phi_0) = \sum_{BPS\ states} e^{2\pi\phi_0 P} \quad (8.2)$$

To illustrate our point we will only consider the contributions from two kinds of states. The exact computation including all possible configurations with total winding W can be done if necessary.

- Consider the configuration of a single multi-wound long string ($w = W$). As we see from (8.1) the momentum p can take the values $p = 0, -1, -2, \dots$ but not $p = 1$ if W is

²⁰ Since we are interested in BPS configurations we have to take all the charge vectors (p_i, w_i) of the various strings to be parallel.

large and positive, since $N_L \geq 0$. For large W the partition function can be computed using the saddle point method²¹ and we find:

$$Z_W^{long} \sim e^{2\pi \frac{1}{\phi_0} W} \quad (8.3)$$

the electric charge for this configuration at the saddle point is :

$$P^{long} = -\frac{W}{\phi_0^2} \quad (8.4)$$

• Now consider the configuration of W singly-wound strings, so for each of them $w = 1$. The momentum of each string can take the values $p = 1, 0, -1, -2, \dots$. We will consider only the contribution from the "ground state"²² $p = 1$. Since we have W strings $P_{total} = W$. The degeneracy for this state is 1 since (8.1) gives $N_L = 0$ and, as we discuss below, we believe that because of the BPS condition there is no contribution to the entropy from the transverse volume²³. This means that the contribution of W singly wound ($w = 1, p = 1$) strings to the partition function is:

$$Z_W^{short} \sim e^{2\pi \phi_0 W} \quad (8.5)$$

and the electric charge:

$$P^{short} = W \quad (8.6)$$

So the partition function for total winding charge W and chemical potential ϕ_0 for momentum will be:

$$Z_W(\phi_0) = Z_W^{short} + Z_W^{long} + \dots \quad (8.7)$$

or

$$Z_W(\phi_0) \sim \exp(2\pi \phi_0 W) + \exp\left(2\pi \frac{1}{\phi_0} W\right) + \dots \quad (8.8)$$

where we have not written the contributions from other configurations. But just from (8.8) we see that in the large W limit, the system will undergo a first order phase transition at

²¹ The degeneracy of a single-string state with momentum p and winding w grows like $\exp(4\pi\sqrt{|pw|})$

²² In the large W limit, the contribution from other states is suppressed.

²³ Except for the center of mass volume factor that is common in both phases and which we do not need to include when we compare the two configurations.

the value of the chemical potential $\phi_0 = 1$. We believe that the exact counting and the inclusion of the other states will not change this result qualitatively.

One might worry that in the multi-string phase we have to include the entropy factors from the volume transverse to the S^1 . We think that since we are interested in a BPS configuration there is only one allowed state of the W short heterotic strings. To understand this better we ask how our configurations look in the supergravity approximation. Clearly the one-long-string looks like a small BPS heterotic black hole. Classically the horizon area is zero but as discussed in [24], [13], [25] we expect that stringy corrections will generate a horizon. Computing the entropy from supergravity with corrections gives the same answer with the microscopic counting. On the other hand for the many-short-string phase we do not have a black hole solution, since in this configuration each of the W heterotic strings has unit winding and unit momentum, so from (8.1) the oscillator level for the left side is $N_L = 0$. The usual BPS heterotic small black holes correspond to winding and momentum charges (W, P) , with W and P large and $PW < 0$, so $N_L \geq 1$. However, there are BPS solutions of the low energy supergravity [26], [27], [28] with the desired charges (W, W) , so now $PW = W^2 > 0$, corresponding to a collection of BPS heterotic strings with $N_L = 0$. Naively these solutions have timelike naked singularities but as explained in [29] they are resolved by the enhançon mechanism. The resulting geometry is regular with no horizon. It is BPS and carries the same charges with the collection of the many singly-wound heterotic strings. So we find it plausible that the many-short-string-phase is described in supergravity by the "heterotic enhançon" solution [30]. This is consistent with our assumption that the number of states in this phase is 1 giving zero entropy, since this solution does not have a horizon²⁴. We think it is interesting to investigate this configuration in possible connections with multi-centered black holes, see also [31], [32], [5], [33].

To summarize we find a first order "phase transition" between a configuration of many singly-wound strings and a single long multi-wound string. Both configurations are BPS. In analogy with the $\mathcal{N} = 4$ gauge theory that we studied before, it is a transition between

²⁴ We could also consider the zero string coupling limit and place the strings in a box of volume V (if we could somehow ignore the issue of having nonzero net charge in compact volume). Then we see that there is only one BPS state: the one where all strings are completely delocalized in the box. Trying to localize any of them would introduce additional 'zero-point' energy and lift the energy of the state above the BPS bound, so naively no volume factors have to be included in the entropy. The phase transition point is independent of the volume V , which we can then send to infinity.

a phase with low energy and zero entropy (short strings) to a phase with large energy and large entropy (long string). We notice from (8.4) and (8.6) that the value of the electric charge P is different in the two phases. In supergravity the two configurations correspond to a smooth solution without a horizon (short strings) and to a small BPS black hole (long string).

We find it interesting that the partition function we considered has an OSV form, since it is computed at fixed magnetic charge W and with chemical potential ϕ_0 for electric charge. One might expect that the OSV index defined on a boundary CFT will receive contributions from both configurations and exhibit the phase transition mentioned above.

9. $D4/D0$ system

We briefly consider the dual system in type IIA, which is a number of $D4$ branes wrapped on $K3$ with $D0$ branes on them. It is well known that a $D4$ -brane wrapped on $K3$ acquires an induced $D0$ -brane charge equal to -1 . So, for N $D4$ and k $D0$ -branes the total $D0$ charge is $Q = k - N$. We want to compute the partition function of BPS states with fixed $D4$ -brane charge N and chemical potential ϕ_0 for $D0$ -brane charge Q :

$$Z_N(\phi_0) = \sum_{BPS\ states} \exp(-2\pi\phi_0 Q) \quad (9.1)$$

• One kind of configurations is that of N $D4$ and $k > 0$ $D0$ branes forming a bound state that looks like a small BPS black hole in type IIA. The contribution from these configurations is:

$$Z_{BH}(\phi_0) = \sum_k d_{N,k} \exp(-2\pi\phi_0(k - N)) \quad (9.2)$$

where $d_{N,k}$ is the number of bound states of N $D4$ -branes and k $D0$ -branes. But we know [18] that for large N and k we have $d_{N,k} \sim \exp(4\pi\sqrt{Nk - N^2})$. Following the analysis of the previous sections, we compute the partition function in the large N limit by a saddle point approximation:

$$Z_{BH}(\phi_0) \sim \exp\left(\frac{2\pi}{\phi_0} N\right) \quad (9.3)$$

Also on the saddle point we find that the $D0$ brane charge for the bound system is equal to :

$$Q_{BH} = \frac{N}{\phi_0^2} \quad (9.4)$$

Some of it is due to the induced $D0$ -brane charge on the $D4$ -branes, so the number of $D0$ -branes is:

$$k = \frac{N}{\phi_0^2} + N \quad (9.5)$$

• The second configuration is that of N $D4$ -branes, with no $D0$ -brane. In this case the corresponding supergravity solution is the enhançon geometry [29], which is smooth and has no horizon. Thus the entropy is zero, or there is only one state. The only $D0$ -brane charge is the induced one, so:

$$Q_{enh} = -N \quad (9.6)$$

And we have:

$$Z_{enh}(\phi_0) \sim \exp(2\pi\phi_0 N) \quad (9.7)$$

Comparing (9.3) and (9.7) we see that in the large N limit there is a phase transition at $\phi_0 = 1$ between the two configurations. In the supergravity description it is a first order transition between an extremal black hole and the enhançon solution for the $D4/D0$ system. This result in IIA is in agreement with the S -dual system of the heterotic string that we studied in the previous section. Again we notice that the two saddle points have the same value of $D4$ -brane charge but different values of $D0$ -brane charge.

10. Holographic Duals

The phase transition in $\mathcal{N} = 4$ on $K3$ that we studied may be related to string theory in yet another way. As for $\mathcal{N} = 4$ on flat space or on $S^3 \times R$, in the large N limit we expect some kind of dual string theory, like in the usual AdS/CFT correspondence. If such a dual string theory exists for $\mathcal{N} = 4$ on $K3$ then we can tell that it will have a phase transition in the large N fixed g_{YM} limit, as a function of g_{YM} . Of course we do not know the exact correspondence between the parameters of $\mathcal{N} = 4$ on $K3$ and its holographic dual string theory, but if we assume that it is similar to the standard AdS/CFT we conclude that this phase transition on the string side will take place in the limit of locally flat space ($N \rightarrow \infty$) and finite string coupling ($g_s \sim g_{YM}^2 \sim const$). The phase transition is with respect to the string coupling.

We also observed that the large N free energy was exactly invariant under S -duality. Let us consider for a moment the $\mathcal{N} = 4$ theory on any other manifold, and let's look at its partition function, possibly with supersymmetry breaking boundary conditions (thermal

partition function). It is reasonable to assume that in the large N , fixed g limit the partition function will have the form:

$$Z_N(g) \sim \exp(N^a f(g)) + \dots \quad (10.1)$$

for some a . Now if we assume that f is exactly invariant under S -duality, as it happened in our system, we have:

$$f\left(\frac{4\pi}{g}\right) = f(g) \quad (10.2)$$

We conclude that the first derivative is either discontinuous at the self-dual point $g = \sqrt{4\pi}$ or exactly zero $f'(\sqrt{4\pi}) = 0$. In the system we studied we found that the first possibility (discontinuity) is realized. Even though we have no indication for it, it would be rather fascinating if the same thing happened for example in $\mathcal{N} = 4$ on $S^3 \times S^1$ in the large N fixed g limit²⁵. Such a possibility, if true, would imply a phase transition for *IIB* string theory in the flat space, finite string coupling limit.

11. Conclusions

We have demonstrated that $\mathcal{N} = 4$ and $\mathcal{N} = 1^*$ on $K3$ exhibit first order phase transitions with respect to the gauge coupling g , in the large N fixed g limit.

There are several aspects of these phase transitions that we find interesting:

- i. They take place in the large N , fixed g limit instead of the usual 't Hooft limit. In the standard AdS/CFT this corresponds to the limit of (locally) flat space and finite string coupling. One then expects that if we knew a holographic dual string theory of $\mathcal{N} = 4$ on $K3$ we would have a phase transition in the string theory at finite string coupling.
- ii. They are phase transitions affecting the supersymmetric partition function, which is usually considered a protected quantity that does not change (or varies smoothly) with respect to continuous changes of the parameters of the theory.
- iii. The same phase transition may exist for the topological or the physical $\mathcal{N} = 4$ on other manifolds. Of course it would be really exciting if such a phase transition existed in

²⁵ If the theory is defined on $M = S^4$ or $M = S^3 \times S^1$, the $SU(N)$ and $SU(N)/\mathbf{Z}_N$ gauge groups give essentially the same results, as $H^2(M, \mathbf{Z}_N) = 0$ for these manifolds and there are no magnetic flux sectors. This indicates that (10.2) must be true in these cases.

the physical $\mathcal{N} = 4$ at finite temperature on S^3 or R^3 as this would imply the same for type IIB string theory at finite string coupling on flat space.

- iv. It would be interesting to study more carefully the interpretation and implications of the analogous phase transitions in the systems of five-branes, heterotic strings etc mentioned in the text and also to investigate possible connections with black hole entropy.
- v. Another point is to study the $\mathcal{N} = 4$ in the large N , fixed g limit on other manifolds or with other gauge groups, for which we have the exact answer from the topologically twisted version. This would provide us with some hint about whether this phase transition is more general or specific to the $K3$ case.

Finally, and in a rather different direction, we think it would be very interesting to study the 't Hooft large N limit of the partition function of the $\mathcal{N} = 4$ theory on $K3$. Presumably there must be some holographic string dual. Since the gauge theory on $K3$ can be twisted to a topological theory, one would expect the same thing from the dual string theory. This has been discussed in [34], [6], [35], [36], [37]. If this dual string theory can be twisted to a "topological" theory, we might be able to compute its partition function exactly, as it was possible for the gauge theory. If all of the above were done then one would have an exact expression for the partition function on both sides of an AdS/CFT-like duality and one could test it exactly in λ or even in N , in analogy with [38] but for a 4-dimensional gauge theory.

12. Acknowledgements

I would like to thank M. Ernebjerg, D. Gaiotto, L. Grant, M. Guica, N. Iizuka, D. Jafferis, S. Lahiri, J. Marsano, J. Minahan, L. Motl, G. Pastras, S. Raju, A. Simons, A. Strominger, C. Vafa for very useful discussions. I would especially like to thank Shiraz Minwalla for numerous invaluable comments, discussions and insights. The work of K.P. was supported in part by DOE grant DE-FG01-91ER40654.

Appendix A. Modular forms

Definitions

The Dedekind eta function $\eta(\tau)$ is defined by:

$$\eta(\tau) = q^{\frac{1}{24}} \prod_{m=1}^{\infty} (1 - q^m), \quad q = \exp(2\pi i\tau). \quad (\text{A.1})$$

It is analytic in the upper half-plane. Under $SL(2, Z)$ it transforms like:

$$\eta(\tau + 1) = e^{i\frac{\pi}{12}} \eta(\tau), \quad \eta\left(-\frac{1}{\tau}\right) = (-i\tau)^{\frac{1}{2}} \eta(\tau). \quad (\text{A.2})$$

We define:

$$G(\tau) = \eta(\tau)^{-24} \quad (\text{A.3})$$

It transforms as:

$$G(\tau + 1) = G(\tau), \quad G\left(-\frac{1}{\tau}\right) = (\tau)^{-12} G(\tau). \quad (\text{A.4})$$

Asymptotic Expansion

The function $G(\tau)$ can be written in the form:

$$G(\tau) = \sum_{n=-1}^{\infty} d_n q^n, \quad q = \exp(2\pi i\tau) \quad (\text{A.5})$$

where d_n are integers. Each d_n is equal to the degeneracy of the level $n+1$ oscillator space of one side of the bosonic string.

We have that $d_{-1} = 1$, $d_0 = 24$ so for large positive values of $Im(\tau)$ the function G has the expansion:

$$G(\tau) = \frac{1}{q} + 24 + \mathcal{O}(q) = e^{-2\pi i\tau} + 24 + \mathcal{O}(e^{2\pi i\tau}) \quad (\text{A.6})$$

Also it is easy to show that for large n we have a Hagedorn-like growth for d_n :

$$d_n \sim \exp(4\pi\sqrt{n}) \quad (\text{A.7})$$

Appendix B. Partition function of $\mathcal{N} = 4$ on $K3$ at large N

In this appendix we will study the large N limit of the partition function (3.3), for τ in the first fundamental domain \mathcal{F}_o of $SL(2, Z)$. Our goal is to identify the largest term in the sum, let us call this term F , as well as the second largest, call it S . We then show that the ratio of F/S grows exponentially in the large N limit. Since the number of terms in

the partition function grows at most like some power of N , we conclude that F contributes exponentially more than all the other terms together, and thus it is reliable to study the large N , fixed τ limit of the partition function by looking only at this largest term F .

Partition function for $\theta = 0$ In the next subsection we present a shorter argument that can prove our result even for $\theta = 0$. In this subsection we develop a longer but more explicit proof for the case $\theta = 0$.

For $\theta = 0$, the coupling $\tau = i\frac{4\pi}{g^2}$ is purely imaginary. We now show that for any $a, b \in \mathbb{R}$ and $a > 0$:

$$|G(ia)| \geq |G(ia + b)| \quad (\text{B.1})$$

From the definition of G :

$$G(\tau) = \frac{1}{q} \prod_{m=1}^{\infty} (1 - q^m)^{-24}, \quad q = \exp(2\pi i\tau) \quad (\text{B.2})$$

we have:

$$\left| \frac{G(ia)}{G(ia + b)} \right| = \prod_{m=1}^{\infty} \left| \frac{1 - e^{2\pi i b m} e^{-2\pi a m}}{1 - e^{-2\pi a m}} \right|^{24} \geq 1 \quad (\text{B.3})$$

since every factor in the product is independently greater than or equal to one.

So in the partition function:

$$Z(\tau) = \frac{1}{N^3} \sum_{0 \leq a, b, d \in \mathbb{Z}; ad=N; b \leq d-1} d G\left(\frac{a\tau + b}{d}\right). \quad (\text{B.4})$$

for the case $\theta = 0$, it is true that for given a and d , the term that will contribute the most is the one with $b = 0$. So, we only need to compare the terms $G\left(\frac{a}{d}\tau\right)$ for $a, d \in \mathbb{N}$ and $ad = N$. The possible values of a and d are the divisors of N . Let us assume that N is even and look at the asymptotic expansion of the various terms for large N , for imaginary $\tau = i\tau_2$:

$$\begin{array}{lll} a = N, d = 1 : & G(N\tau) \sim & \exp(N2\pi\tau_2) \\ a = N/2, d = 2 : & G(N/4\tau) \sim & \exp\left(\frac{N}{4}2\pi\tau_2\right) \\ \dots & \dots & \dots \\ a = 2, d = N/2 : & G\left(\frac{4}{N}\tau\right) \sim G\left(-\frac{N}{4}\frac{1}{\tau}\right) \sim & \exp\left(\frac{N}{4}2\pi\frac{1}{\tau_2}\right) \\ a = 1, d = N : & G\left(\frac{1}{N}\tau\right) \sim G\left(-N\frac{1}{\tau}\right) \sim & \exp\left(N2\pi\frac{1}{\tau_2}\right) \end{array} \quad (\text{B.5})$$

It is clear that it is either the term $G(N\tau)$ or $G\left(\frac{\tau}{N}\right)$ that will be the largest for large enough N . Specifically in the first fundamental domain \mathcal{F}_o and for $\theta = 0$ we have $\tau_2 > 1$ so $G(N\tau)$ is the largest. We also see that the second largest term (no matter which one it is) is exponentially suppressed compared to $G(N\tau)$ (for $\tau_2 > 1$). Similarly for $\tau_2 < 1$ the dominant term is $G\left(\frac{\tau}{N}\right)$. It is clear that the same will happen with any other divisor of N . This completes our analysis for $\theta = 0$.

Partition function for $\theta \neq 0$

Now we claim that for τ complex and in the first fundamental domain \mathcal{F}_o of $SL(2, Z)$ it is still $G(N\tau)$ that is the largest term.

As we take $N \rightarrow \infty$ there are some terms in the partition function, for which the argument of the function G has the largest imaginary part, and which goes to infinity $N \rightarrow \infty$. These are the terms for which $a > d$. From them the term $G(N\tau)$ is clearly the dominant one. There are also terms that have the smallest imaginary part, which goes to zero as $N \rightarrow \infty$. These are the terms for which $a < d$. For them we do a modular transformation and then they get a very large and positive imaginary part that goes like: $\frac{Na\tau_2}{a^2\tau_2^2 + (a\tau_1 + b)^2}$. It is easy to show that in the first fundamental domain the largest of these is the one with $a = 1$, $b = 0$, which corresponds to the term $G\left(\frac{\tau}{N}\right)$. In \mathcal{F}_o we can see that $G(N\tau)$ grows faster with N .

So again we conclude that the largest term in \mathcal{F}_o is $G(N\tau)$ and the second-largest term is exponentially smaller.

Appendix C. Instanton moduli spaces on $K3$

In this appendix we will try to find the behavior of the Euler characteristic $c_{k,N}$ of the k -instanton moduli space for $SU(N)$ on $K3$, for large N and k . Let's assume that N is prime for simplicity.

The partition function can be written in two different forms. One is the general form where the contribution from the various instanton sectors are grouped together:

$$Z_N(\tau) \sim q^{-N} \sum_{k=0}^{\infty} c_{k,N} q^k, \quad q = \exp(2\pi i \tau) \quad (\text{C.1})$$

and the other is the function that was computed in [1]:

$$Z_N(\tau) = \frac{1}{N^3} G(N\tau) + \frac{1}{N^2} \sum_{m=0}^{N-1} G\left(\frac{\tau + m}{N}\right), \quad (\text{C.2})$$

where

$$G(\tau) = \eta(\tau)^{-24} \quad (\text{C.3})$$

Now if we want to compute $c_{k,N}$, according to (C.1) we have to isolate the coefficient of q^{k-N} in (C.2).

- First let's study the sum over m in (C.2). We will use the expansion (A.5) for $G(\tau)$ that we mentioned in Appendix A:

$$\sum_{m=0}^{N-1} G\left(\frac{\tau+m}{N}\right) = \sum_{k=-1}^{\infty} d_k q^{\frac{k}{N}} \sum_{m=0}^{N-1} \exp\left(2\pi i \frac{mk}{N}\right) \quad (\text{C.4})$$

The sum over m forces k to be a multiple of N so we have:

$$\sum_{m=0}^{N-1} G\left(\frac{\tau+m}{N}\right) = N \sum_{l=0}^{\infty} d_{lN} q^l \quad (\text{C.5})$$

For $c_{k,N}$ we want the coefficient of q^{k-N} . It is equal to d_{kN-N^2} (up to multiplicative factors that we ignore). Now using (A.7) we find:

$$c_{k,N} \sim \exp\left(4\pi\sqrt{kN - N^2}\right) \quad (\text{C.6})$$

- Now let's look at the term $G(N\tau)$. We have:

$$G(N\tau) = \sum_{n=-1}^{\infty} d_n q^{Nn} \quad (\text{C.7})$$

This term can contribute to the k -instanton sector only if k is a multiple of N but even then its contribution is $d_{\frac{k}{N}-1}$ which is suppressed compared to (C.6), so we can ignore it in the large N limit.

Thus our final result for the Euler characteristic of the k -instanton moduli space for $SU(N)$ on $K3$ is (C.6).

References

- [1] C. Vafa and E. Witten, “A Strong coupling test of S duality,” Nucl. Phys. B **431**, 3 (1994) [arXiv:hep-th/9408074].
- [2] J. A. Minahan, D. Nemeschansky, C. Vafa and N. P. Warner, “E-strings and $N = 4$ topological Yang-Mills theories,” Nucl. Phys. B **527**, 581 (1998) [arXiv:hep-th/9802168].
- [3] J. M. F. Labastida and C. Lozano, “The Vafa-Witten theory for gauge group $SU(N)$,” Adv. Theor. Math. Phys. **3**, 1201 (1999) [arXiv:hep-th/9903172].
- [4] C. Montonen and D. Olive, Phys. Lett. B **72** (1977) 117; P. Goddard, J. Nyuts and D. Olive, Nucl. Phys. B **125** (1977) 1.
- [5] H. Ooguri, A. Strominger and C. Vafa, “Black hole attractors and the topological string,” Phys. Rev. D **70**, 106007 (2004) [arXiv:hep-th/0405146].
- [6] C. Lozano, “Duality in topological quantum field theories,” arXiv:hep-th/9907123.
- [7] J. P. Yamron, “Topological Actions From Twisted Supersymmetric Theories,” Phys. Lett. B **213**, 325 (1988).
- [8] N. Marcus, “The Other topological twisting of $N=4$ Yang-Mills,” Nucl. Phys. B **452**, 331 (1995) [arXiv:hep-th/9506002].
- [9] E. Witten, “Supersymmetric Yang-Mills theory on a four manifold,” J. Math. Phys. **35**, 5101 (1994) [arXiv:hep-th/9403195].
- [10] E. Witten, “Topological Quantum Field Theory,” Commun. Math. Phys. **117**, 353 (1988).
- [11] A. Dabholkar and J. A. Harvey, “Nonrenormalization Of The Superstring Tension,” Phys. Rev. Lett. **63**, 478 (1989).
- [12] A. Dabholkar, G. W. Gibbons, J. A. Harvey and F. Ruiz Ruiz, “Superstrings And Solitons,” Nucl. Phys. B **340**, 33 (1990).
- [13] A. Dabholkar, “Exact counting of black hole microstates,” Phys. Rev. Lett. **94**, 241301 (2005) [arXiv:hep-th/0409148].
- [14] R. Dijkgraaf, E. P. Verlinde and M. Vonk, “On the partition sum of the NS five-brane,” arXiv:hep-th/0205281.
- [15] L. Girardello, A. Giveon, M. Porrati and A. Zaffaroni, “S duality in $N=4$ Yang-Mills theories with general gauge groups,” Nucl. Phys. B **448**, 127 (1995) [arXiv:hep-th/9502057].
- [16] L. Girardello, A. Giveon, M. Porrati and A. Zaffaroni, “NonAbelian strong - weak coupling duality in (string derived) $N=4$ supersymmetric Yang-Mills theories,” Phys. Lett. B **334**, 331 (1994) [arXiv:hep-th/9406128].
- [17] D. J. Gross and E. Witten, “Possible Third Order Phase Transition In The Large N Lattice Gauge Theory,” Phys. Rev. D **21**, 446 (1980).

- [18] C. Vafa, “Instantons on D-branes,” Nucl. Phys. B **463**, 435 (1996) [arXiv:hep-th/9512078].
- [19] C. Vafa, “Gas of D-Branes and Hagedorn Density of BPS States,” Nucl. Phys. B **463**, 415 (1996) [arXiv:hep-th/9511088].
- [20] J. Polchinski and M. J. Strassler, “The string dual of a confining four-dimensional gauge theory,” arXiv:hep-th/0003136.
- [21] R. Dijkgraaf and C. Vafa, “A perturbative window into non-perturbative physics,” arXiv:hep-th/0208048.
- [22] J. M. Maldacena and A. Strominger, “AdS(3) black holes and a stringy exclusion principle,” JHEP **9812**, 005 (1998) [arXiv:hep-th/9804085].
- [23] R. Dijkgraaf, J. M. Maldacena, G. W. Moore and E. P. Verlinde, “A black hole farey tail,” arXiv:hep-th/0005003.
- [24] A. Sen, “Extremal black holes and elementary string states,” Mod. Phys. Lett. A **10**, 2081 (1995) [arXiv:hep-th/9504147].
- [25] A. Dabholkar, R. Kallosh and A. Maloney, “A stringy cloak for a classical singularity,” JHEP **0412**, 059 (2004) [arXiv:hep-th/0410076].
- [26] K. Behrndt, “About a class of exact string backgrounds,” Nucl. Phys. B **455**, 188 (1995) [arXiv:hep-th/9506106].
- [27] R. Kallosh and A. D. Linde, “Exact supersymmetric massive and massless white holes,” Phys. Rev. D **52**, 7137 (1995) [arXiv:hep-th/9507022].
- [28] M. Cvetič and D. Youm, “Singular BPS saturated states and enhanced symmetries of four-dimensional N=4 supersymmetric string vacua,” Phys. Lett. B **359**, 87 (1995) [arXiv:hep-th/9507160].
- [29] C. V. Johnson, A. W. Peet and J. Polchinski, “Gauge theory and the excision of repulson singularities,” Phys. Rev. D **61**, 086001 (2000) [arXiv:hep-th/9911161].
- [30] M. Natsuume, “The heterotic enhancon,” Phys. Rev. D **65**, 086002 (2002) [arXiv:hep-th/0111044].
- [31] J. M. Maldacena, J. Michelson and A. Strominger, “Anti-de Sitter fragmentation,” JHEP **9902**, 011 (1999) [arXiv:hep-th/9812073].
- [32] F. Denef, “Quantum quivers and Hall/hole halos,” JHEP **0210**, 023 (2002) [arXiv:hep-th/0206072].
- [33] R. Dijkgraaf, R. Gopakumar, H. Ooguri and C. Vafa, “Baby universes in string theory,” arXiv:hep-th/0504221.
- [34] C. M. Hull, “Timelike T-duality, de Sitter space, large N gauge theories and topological field theory,” JHEP **9807**, 021 (1998) [arXiv:hep-th/9806146].
- [35] P. de Medeiros, C. M. Hull, B. J. Spence and J. M. Figueroa-O’Farrill, “Conformal topological Yang-Mills theory and de Sitter holography,” JHEP **0208**, 055 (2002) [arXiv:hep-th/0111190].

- [36] A. Neitzke and C. Vafa, “N = 2 strings and the twistorial Calabi-Yau,” arXiv:hep-th/0402128.
- [37] R. Dijkgraaf, S. Gukov, A. Neitzke and C. Vafa, “Topological M-theory as unification of form theories of gravity,” arXiv:hep-th/0411073.
- [38] R. Gopakumar and C. Vafa, “On the gauge theory/geometry correspondence,” Adv. Theor. Math. Phys. **3**, 1415 (1999) [arXiv:hep-th/9811131].